Mini Project report on

OCR using Neural Networks

Submitted by

Madhusoodan M Pataki 2SD14CS054

Anarghya K Jantali 2SD14CS128

(6th sem students)

Under the guidance of

Prof. Sharada H. N.

Abstract of project

Optical Character Recognition (OCR) deals with recognition and conversion of printed characters in a scanned document that is taken as an input to an editable, digitized text. This improves the storage, processing, mining that are pivotal in further development of essential applications that serve various domains. We are implementing such a process, which takes a scanned/captured image as an input, pre-process the image, recognize the characters of English language and output in a form of string to a text document. Current scope is limited to English characters and digits. Recognition is also limited to printed characters rather than manuscript/handwriting. This process can be improvised to recognize various languages as well as handwritten materials. The input image is put through pre-processing, segmentation and feature extraction. Feature extraction and recognition of characters is implemented with Neural-Network. Using neural network to train and recognize the characters increases the coverage as well as accuracy of the results compared to conventional horizontal and vertical feature extraction methods. With increased accuracy, this system can be implemented in applications such as form reading such as cheque, postal code, official letters etc. This can also be augmented with a dictionary and intelligent interpreter to evaluate answer scripts etc.

Department of Computer Science & Engineering

S.D.M. College of Engineering and Technology,

Dharwad-02

2016-17